



IB/CS 385 INTRODUCTION TO BUSINESS PROGRAMMING AND ARTIFICIAL INTELLIGENCE

IES Abroad Madrid

DESCRIPTION:

This course enables students to acquire basic skills in one of the most fundamental skills of the present period: computer programming. The course introduces students to data analysis programs with the aim of growing their skills in using those programs to predict future outcomes. Students will first learn how to write programming codes from scratch using Python, one of the most intuitive and popular programming languages. The course will then familiarize students with the most important predictive techniques – from linear regressions to machine learning – and ask them to implement what they’ve learned theoretically in applications related to management and marketing.

CREDITS: 3 credits

CONTACT HOURS: 45 hours

LANGUAGE OF INSTRUCTION: English

INSTRUCTOR: Instructor Name

PREREQUISITES: An introductory knowledge of statistics is required. Students must bring their own laptop in class”

ADDITIONAL COST: None

METHOD OF PRESENTATION:

Theoretical lectures illustrated by examples followed by programming sessions. Students should plan to bring their own laptops (at least 2 GB of RAM) to class on the dates indicated in the schedule. In the first session all relevant concepts in calculus and statistics will be introduced to students.

Since the conceptual material is dense and requires both theoretical and practical modules to be adequately processed and understood by students, the course often pairs theoretical sessions with lab components during which students will get hands-on practice with the week's material by working on assigned lab activities

REQUIRED WORK AND FORM OF ASSESSMENT:

- Class Engagement - 15%
- Midterm Exam - 30%
- Group Project- 25%
- Final Exam - 30%

Class Engagement

Students are expected to actively participate in class and to demonstrate respect and responsibility in the classroom. A rubric for participation will be available on Moodle and distributed on the first day of class. During practical class sessions, students will be expected to engage actively with self-assessment tests and practical exercises requiring them to code and resolve business problems.

Midterm Exam

Students will be asked to resolve small exercises and to answer multiple-choice questions on the techniques covered in the first half of the class. The exam will last 1 hour.

Group Project

Students will be asked to complete a group project, using coding to answer a series of business questions.

Final Exam Students will be asked to resolve small exercises and to answer multiple-choice questions on the techniques covered in the first half of the class. The exam will last 1 hour.

LEARNING OUTCOMES:

By the end of the course, students will be able to:

- Apply the logic of computer programming and Artificial Intelligence
- Write code that collects, manipulates, and visualizes data in Python
- Evaluate a data set and identify analytical techniques that can extract relevant information from it
- Implement the most important predictive techniques in Python and apply them to financial or business concerns

ATTENDANCE POLICY:

As a member of our class community, you are expected to be present and on time every day. Attending class has an impact on your learning and academic success. For this reason, attendance is mandatory for all IES Madrid classes, including course-related excursions. Absences will only be excused in cases of documented medical or family emergencies. You are allowed a maximum of one unexcused absence for once-a-week courses and two unexcused absences for twice-a-week courses. Each additional unexcused absence will result in a deduction of 3 percentage points from your final grade and/or disciplinary action. Accumulating 7 or more unexcused absences in twice-a-week courses, or 4 or more in once-a-week courses, may result in a failing grade for the course.

Late or missed submission of assessed work: any assessed work, including exams, tests, and presentations, will be rescheduled only if the absence is justified; otherwise, it will be graded as zero.

Punctuality is critical: three instances of being over 10 minutes late will be counted as one absence and being more than 15 minutes late will also count as one absence.

On museum days, students must arrive exactly at the scheduled entrance time. If a student arrives late, they are responsible for entering the museum independently by presenting their student ID and passport to verify their identity as a student at the entrance and purchasing their own ticket. Once inside, they must consult the syllabus to identify their assigned class and locate their group. Students are expected to take responsibility for locating the class, this includes reviewing the syllabus, Moodle and class information, and considering having alternative means of contact. In such cases, the student will be marked as late. Failure to join their group after arriving late will be recorded as an absence.

For more information, please refer to the full Attendance and Punctuality Policy document.

CONTENT:

Sessions	Content	Assignments
Session 1	• Course Introduction • Python initial assessment (not graded)	• Professor's Slide • Initial Assessment
Session 2	Introduction to statistical learning and model accuracy. Brief history of AI Defining data science, machine learning, deep learning, artificial intelligence, large language models Why AI is at a tipping point (cloud, open source, data explosion, models improvements)	Readings 1. Introduction to Machine Learning: https://towardsdatascience.com/introduction-to-machinelearning-for-beginners-eed6024fdb08 Cheatsheets
Session 3	Statistical learning and model metrics. Classification of Analytics: descriptive, diagnostic, predictive analytics Data Science Process (CRISP Model) Model performances: overfitting, underfitting Introduction to Google Colab	Readings 1. Types of Analytics: https://business.adobe.com/blog/basics/descriptivepredictiveprescriptive-analytics-explained 2. CRISP DM Model: https://www.datascience-pm.com/crisp-dm-2/Cheatsheets

Sessions	Content	Assignments
Session 4	Lab session: introduction to Python (<u>bring your own laptop</u>) Introduction to Python: basic commands, data structures, if-then-else, for, while cycles, functions	Notebooks Python notebooks in Intro Folder Cheatsheets
Session 5	Data Exploration and Visualization: univariate and multivariate statistics Problem / Objective: visualization (barplot, histogram, scatterplot, time series, boxplot, correlation matrices, etc.) Main statistics: mean, median, mode, variance, standard deviation, count, correlation, etc.	Notebooks Python notebooks in Intro Folder Cheatsheets Readings 1. EVALUATION METRICS https://www.analyticsvidhya.com/blog/2019/08/11-important-modevaluation-error-metrics/ 2. UNDERFITTING VS OVERFITTING: https://towardsdatascience.com/overfitting-and-underfittingprinciplesea8964d9c45c
Session 6	Data Exploration and Visualization: univariate and multivariate statistics II Problem / Objective: visualization (barplot, histogram, scatterplot, time series, boxplot, correlation matrices, etc.) Main statistics: mean, median, mode, variance, standard deviatio	Notebooks Python notebooks in EDA Folder Cheatsheets Readings 1. EDA (Exploratory Data Analysis) https://towardsdatascience.com/exploratory-dataanalysis8fc1cb20fd15 2. CLEANING DATA (with PANDAS and NUMPY): https://realpython.com/python-data-cleaning-numpy-pandas/
Session 7	Linear regression part 1	Readings 1. Linear Regression Intoduction: A Beginner's Guide to Linear Regression in Python with Scikit-Learn - KDnuggets Cheatsheets
Session 8	Lab session: linear regression	Notebooks Python notebooks in Linear Regression Folder
Session 9	Linear regression part 2	Readings 1. Linear Regression using Python: https://towardsdatascience.com/simple-linear-regressionmodel-using-python-machine-learning-eab7924d18b4
Session 10	Lab session linear regression part 2	Notebooks Python notebooks in Linear Regression Folder

Sessions	Content	Assignments
Session 11	Classification introduction Main Classification Use Cases Confusion Matrix Performance Metrics	Readings 1. Classification introduction: https://www.datacamp.com/blog/classification-machinelearning Cheatsheets
Session 12	MIDTERM EXAM	Midterm Exam
Session 13	Classification Models Logistic Regression	Notebooks Python notebooks in Classification Folder Cheatsheets Readings 1. Logistic Regression Intro: https://www.analyticsvidhya.com/blog/2021/08/conceptual-understanding-of-logistic-regression-for-data-science-beginners/
Session 14	Classification Models Decision Tree	Notebooks Python notebooks in Classification Folder Cheatsheets Readings 1. Decision Tree: https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html
Session 15	Classification Models Random Forest	Notebooks Python notebooks in Classification Folder Cheatsheets Readings 1. Random Forest: https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/
Session 16	Unsupervised learning Principal Component Analysis 1	Cheatsheets Principal Component Analysis: https://www.sartorius.com/en/knowledge/science-snippets/what-is-principal-component-analysis-pca-and-how-it-is-used-507186
Session 17	Unsupervised learning: Principal Component Analysis 2	Notebooks Python notebooks in PCA and Clustering Folder
Session 18	Unsupervised learning Clustering 1	Cheatsheets 1. Clustering: https://blog.dataiku.com/clustering-how-it-works-inplain-english
Session 19	Unsupervised learning Clustering 2	Notebooks Python notebooks in PCA and Clustering Folder

Sessions	Content	Assignments
Session 20	Data Modeling Lab	Notebooks Exercises in class
Session 21	Data Modeling Lab	Notebooks Exercises in class DEADLINE: GROUP PROJECT SUBMISSION
Session 22	Final review of the course / Exam simulation	Multiple-choice and open-ended questions

REQUIRED READINGS:

- Types of Analytics: <https://business.adobe.com/blog/basics/descriptive-predictive-prescriptive-analytics-explained>
- CRISP DM Model: <https://www.datascience-pm.com/crisp-dm-2/>
- EVALUATION METRICS <https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/>
- UNDERFITTING VS OVERFITTING: <https://towardsdatascience.com/overfitting-and-underfitting-principles-ea8964d9c45c>
- EDA (Exploratory Data Analysis) <https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15>
- CLEANING DATA (with PANDAS and NUMPY): <https://realpython.com/python-data-cleaning-numpy-pandas/>
- Linear Regression Introduction: A Beginner's Guide to Linear Regression in Python with Scikit-Learn - KDnuggets
- Linear Regression using Python: <https://towardsdatascience.com/simple-linear-regression-model-using-python-machinelearning-eab7924d18b4>
- Classification introduction: <https://www.datacamp.com/blog/classification-machine-learning>
- Logistic Regression Intro: <https://www.analyticsvidhya.com/blog/2021/08/conceptual-understanding-of-logistic-regression-for-data-science-beginners/>
- Decision Tree: <https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html>
- Random Forest: <https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/>
- Principal Component Analysis: <https://www.sartorius.com/en/knowledge/science-snippets/what-is-principal-componentanalysis-pca-and-how-it-is-used-507186>
- Clustering: <https://blog.dataiku.com/clustering-how-it-works-in-plain-english>

RECOMMENDED READINGS:

- Short history of AI 1: <https://www.freecodecamp.org/news/the-history-of-ai/>
- Short history of AI 2: <https://www.tableau.com/data-insights/ai/history>
- What is Generative AI?: <https://www.gartner.com/en/insights/generative-ai-for-business>
- Jake Vanderplas, Python Data Science Handbook: Essential Tools for Working With Data, O'Reilly, 2022: <https://www.oreilly.com/library/view/python-data-science/9781098121211/>

INSTRUCTOR BIOGRAPHY:

Brief biography of the faculty.